

**Claims**

What is claimed is:

5        1. A method of synchronizing visual information with audio playback, comprising the steps of:

         receiving a user selection of a desired audio file;

         initiating play of the desired audio file; and

         displaying visual information associated with the  
10 desired audio file on a display device in accordance with timestamp data such that the visual information is displayed synchronously with the playing of the desired audio file, wherein the commencement of playing the desired audio file and the commencement of the displaying step are a function of a signal from the display device.

         2. The method as recited in claim 1 wherein the signal comprises an infrared signal.

20        3. The method as recited in claim 1 wherein the display device comprises a handheld device.

         4. The method as recited in claim 1 wherein the signal turns the remote device off and on prior to the remote device  
25 playing the desired audio file.

         5. The method as recited in claim 1 further comprising the step of analyzing the audio file and generating timestamp data corresponding to the visual information associated with the  
30 audio file.

6. The method as recited in claim 5 wherein the timestamp data is generated utilizing a text based process.

7. The method as recited in claim 5 wherein the timestamp data is generated utilizing a pronunciation dictionary process.

8. The method as recited in claim 5 wherein the timestamp data is generated utilizing a note transcription of music process.

9. The method as recited in claim 5 wherein the timestamp data is generated utilizing tempo information extracted from music.

10. The method as recited in claim 5 further comprising the step of comparing a location of a keyword extracted from the desired audio file with an actual location of a keyword within the timestamp data and adjusting the location of the extracted keyword to match the location of the keyword within the timestamp data.

11. The method as recited in claim 1 wherein the visual information comprises text lines.

12. The method as recited in claim 1 further comprising the step of displaying a title of the desired audio file on the display device.

13. The method as recited in claim 1 wherein the selection is from a list stored in memory associated with the display device.

14. The method as recited in claim 1 wherein the selection is performed by a random number selection method.

15. The method as recited in claim 1 further comprising the step of sending a signal from the display device to a remote device to cause the remote device to start.

16. A method of synchronizing visual information with audio playback, comprising the steps of:

extracting at least one acoustic feature from audio data;

detecting pauses within the audio data;

segmenting the audio data into at least one segment in accordance with the at least one acoustic feature and the pauses;

generating at least one timestamp value associated with the at least one segment;

assigning the at least one timestamp value to the at least one segment such that each segment may be displayed synchronously with the audio playback; and

displaying the at least one segment synchronously with the audio playback.

17. The method as recited in claim 16 wherein the at least one segment refers to lyrics of a song.

18. The method as recited in claim 16 wherein the distinguishing step includes a differentiation between segments intended to be sung by one of the male gender and the female gender.

19. The method as recited in claim 16 further comprising the step of providing an indication of a tempo of the visual information corresponding to the audio playback.

20. The method of synchronizing visual information with audio playback as recited in claim 16 wherein the pause detection step separates the audio data into voice segments and non-voice segments.

21. A method of generating timestamp data, from an audio source comprising the steps of:

extracting voice and non-voice data from said source;  
analyzing the voice and non-voice data to identify selected information in the voice and non-voice data, the selected information providing a basis for generating timestamps; and

generating timestamp values associated with each of the selected information.

22. The method of generating timestamp data as recited in claim 21 wherein the analyzing step includes the step of dividing the voice and non-voice data into separate segments.

23. An apparatus for synchronizing visual information associated with audio playback, comprising:

a feature extraction device for extracting acoustical features from audio data;

a pause detector device for detecting pauses in the audio data;

a classifier device for parsing a continuous bit-stream of audio data into different non-overlapping segments such that each segment is homogenous in terms of its class; and

a timestamp device for assigning timestamp values to each segment.

24. An apparatus for synchronizing visual information associated with audio playback, comprising:

a memory; and

a processor configured to receive a user selection of a desired audio file; initiate play of the desired audio file; and display visual information associated with the desired audio file on a display device in accordance with timestamp data such that the visual information is displayed synchronously with the playing of the desired audio file, wherein the commencement of playing the desired audio file and the commencement of the displaying step are a function of a signal from the display device.